Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 5346–5350, June 1992
Genetics

A 71-kilodalton protein is a major product of the Duchenne muscular dystrophy gene in brain and other nonmuscle tissues

(*DMD* gene products/Dp71/gene expression/alternative splicing)

DORON LEDERFEIN*, ZEHAVA LEVY*, NATALIE AUGIER†, DOMINIQUE MORNET†, GLENN MORRIS‡, ORA FUCHS*, DAVID YAFFE*, AND URI NUDEL*

*Department of Cell Biology, The Weizmann Institute of Science, Rehovot 76100, Israel; †Pathologie Moléculaire du Muscle, Institut National de la Santé et de la Recherche U300, Faculté de Pharmacie, Av. Ch. Flahaut, 34060 Montpellier Cedex 1, France; and ‡The North East Wales Institute, Kelsterton Road, Connah's Quay, Deeside, Clwyd, U.K.

Communicated by Ephraim Katchalski-Katzir, February 19, 1992

ABSTRACT **The known Duchenne muscular dystrophy (DMD) gene products, the muscle- and brain-type dystrophin isoforms, are 427-kDa proteins translated from 14-kilobase (kb) mRNAs. Recently we described a 6.5-kb mRNA that also is transcribed from the *DMD* gene. Cloning and *in vitro* transcription and translation of the entire coding region show that the 6.5-kb mRNA encodes a 70.8-kDa protein that is a major product of the *DMD* gene. It contains the C-terminal and the cysteine-rich domains of dystrophin, seven additional amino acids at the N terminus, and some modifications formed by alternative splicing in the C-terminal domain. It lacks the entire large domain of spectrin-like repeats and the actin-binding N-terminal domain of dystrophin. This protein is the major *DMD* gene product in brain and other nonmuscle tissues but is undetectable in skeletal muscle extracts.**

Duchenne muscular dystrophy (DMD) is a lethal X chromosome-linked recessive disease, characterized by a progressive degeneration of muscles. A significant proportion of DMD patients also suffers from mental retardation (reviewed in ref. 1). The huge gene, which is defective in DMD patients, spans over 2300 kilobases (kb). The transcription product of the gene in the muscle is a 14-kilobase (kb) mRNA encoding a 427-kDa rod-shaped protein called dystrophin, which is associated with the sarcolemma (2–6). The predicted amino acid sequence of dystrophin indicates that it consists of four domains: an N-terminal domain that shares sequence similarity with the actin-binding domain of *Dictyostelium* α-actinin, a spectrin-like domain consisting of 24 repeats of triple-helix structures, a cysteine-rich domain sharing homology with the Ca^{2+}-binding domain of *Dictyostelium* α-actinin, and a unique C-terminal domain (2). A 14-kb mRNA produced by the same gene, which encodes a protein very similar to the muscle-type dystrophin but regulated by a different promoter, is present in the brain (7–14).

Recently some of us identified a 6.5-kb mRNA transcribed from the same gene that seems to be the major *DMD* gene product in many nonmuscle tissues including brain and apparently is regulated by a different promoter(s) (15). Here we report the cloning of the entire coding sequence of the 6.5-kb mRNA and the identification of a protein in Western blots of cell extracts that comigrates with the *in vitro* translation product of the cloned cDNA. This protein, which differs greatly from dystrophin, seems to be a main product of the *DMD* gene; its level in several nonmuscle tissues is comparable to the amount of dystrophin in muscle.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

METHODS

cDNA Cloning, Screening of Libraries, and DNA Sequencing. Screening of cDNA libraries, purification of phages and preparation of DNA, subcloning of DNA fragments into plasmid vectors, and preparation of plasmid DNA were done essentially as described in the manual of Sambrook *et al.* (16). PCR-based 5′ extension of cDNA clones was done as described by Frohman *et al.* (17). Plasmid DNA was sequenced by using the Sequenase kit (United States Biochemical).

Preparation of Tissue and Cell Extracts. Freshly isolated rat tissues (0.4 g) were homogenized in 5 ml of electrophoresis sample buffer [67.5 mM Tris·HCL, pH 6.8/20% (vol/vol) glycerol/15% SDS/5% 2-mercaptoethanol/0.001% bromophenol blue] by using a Polytron homogenizer. Human hepatoma Hep G2 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. The cells were collected in sample buffer and homogenized. All samples were heated for 20 min at 60°C after homogenization.

Western Blot Analysis. Protein samples (30–40 μg) were size-fractionated on 3–10% polyacrylamide/SDS gradient gels as described by Laemmli (18) and modified by Pons *et al.* (19). After electrophoresis for 90 min at 120 mA, the proteins were transferred to nitrocellulose filters; blotting was performed for 18 hr at 150 mA at 4°C. Staining with the first and second antibodies was done with the Protoblot Western blot AP (alkaline phosphatase) system from Promega.

Three antibodies were used in this study. MANDRA1 is a monoclonal antibody (mAb) raised against a fusion protein containing the C-terminal 485 amino acids of human dystrophin (20). The antibody reacts with an epitope within amino acids 3558–3684 of dystrophin. This mAb is specific to dystrophin and does not react with the dystrophin-related protein (21) encoded by a gene located on human chromosome 6. MANDYS1 is a mAb that reacts with an epitope in the spectrin-like domain of dystrophin (22). H-2A12 is a mAb raised against a fusion protein containing amino acids 3357–3660 of chicken dystrophin; this antibody reacts with an epitope within amino acids 3357–3450 of the protein.

RESULTS

Cloning of the Entire Coding Sequence of the 6.5-kb mRNA. RNase protection assays and partial cDNA cloning suggested that the 6.5-kb mRNA shares with dystrophin all or most of the 3′ untranslated region and the regions encoding the C-terminal domain and cysteine-rich domain (15). Sequencing of brain, liver, testis, and hepatoma cDNA clones isolated with probes for these regions revealed a unique sequence§

Abbreviation: DMD, Duchenne muscular dystrophy.
§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M92650).

that is derived from an exon not represented in dystrophin mRNA and located 5′ to the common region encoding the cysteine-rich domain (in the intron between exons 62 and 63; ref. 23). On the basis of sequence divergence between human, rat, and mouse cDNA clones and the existence of an in-frame stop codon in the mouse and rat sequence, we have identified the putative AUG translation initiation codon of the 6.5-kb mRNA. It is located seven codons upstream from the point of divergence between the common and the unique sequence (Fig. 1 *Upper*). The sequence around this AUG fits the Kozak consensus sequence (24).

The expected identity in sequence between substantial parts of the 6.5-kb mRNA and the 14-kb mRNA made it very difficult to conclusively determine the entire coding sequence of the 6.5-kb mRNA from the sequences of partially overlapping small cDNA clones. Therefore, we attempted to clone the entire coding sequence in one clone by PCR with a 5′ primer specific for the 6.5-kb mRNA. First-strand cDNA was synthesized by using RNA from human amniotic fluid cells (these cells express the 6.5-kb mRNA; H. Prigojin, R. Shomrat, U.N. and D.Y., unpublished work) and a primer complementary to a sequence in the 3′ untranslated region of dystrophin mRNA (TGCATAGACGTGTAAAACCTGCC; nucleotides 11519–11541 of human dystrophin cDNA; ref. 2). The cDNA was amplified by PCR with the same 3′ primer and a primer containing a sequence specific to the 6.5-kb mRNA, located 28 nucleotides upstream of the initiator AUG (Fig. 1 *Upper*). The 2.1-kb amplified PCR product was cloned into the plasmid vector Gemini 3, and sequenced (plasmid HMD2.1). The sequence of the region common to the 6.5-kb mRNA and the 14-kb dystrophin mRNA was identical to that of the homologous region in human dystrophin cDNA, except for a T to C substitution (position 11143), which does not change the encoded protein sequence (this substitution may reflect a polymorphism at this position or could have occurred during the amplification of the cDNA by PCR). In addition, two exons of dystrophin mRNA (nucleotides 10432–10470 and 11221–11252) were missing in the cloned cDNA (Fig. 1 *Lower*). Alternative splicing in these two positions has been reported (11, 15). Feener *et al.* (11) attributed this alternative splicing pattern, found in both human skeletal muscle and brain cDNA clones, to dystrophin mRNA. Bar *et al.* (15) identified this pattern in a mouse liver cDNA clone and attributed it to the 6.5-kb mRNA. It is possible that this pattern of splicing, reported by Feener *et al.* (11) originated from the 6.5-kb mRNA and not the 14-kb brain or muscle dystrophin mRNA. It also should be pointed out that our sequencing data of a genomic DNA clone indicate that the exact position of the second alternatively spliced exon in mice is nucleotides 11219–11250; the position indicated by Feener *et al.* (11) is 11221–11252 (for human mRNA). Because of sequence redundancy near the 5′ and the 3′ splice sites, the 2-base-pair (bp) shift in the assignment of the exon borders does not change the sequence of the encoded mRNAs. While the deletion of the first exon (39 bp) does not change the reading frame, the deletion of the second

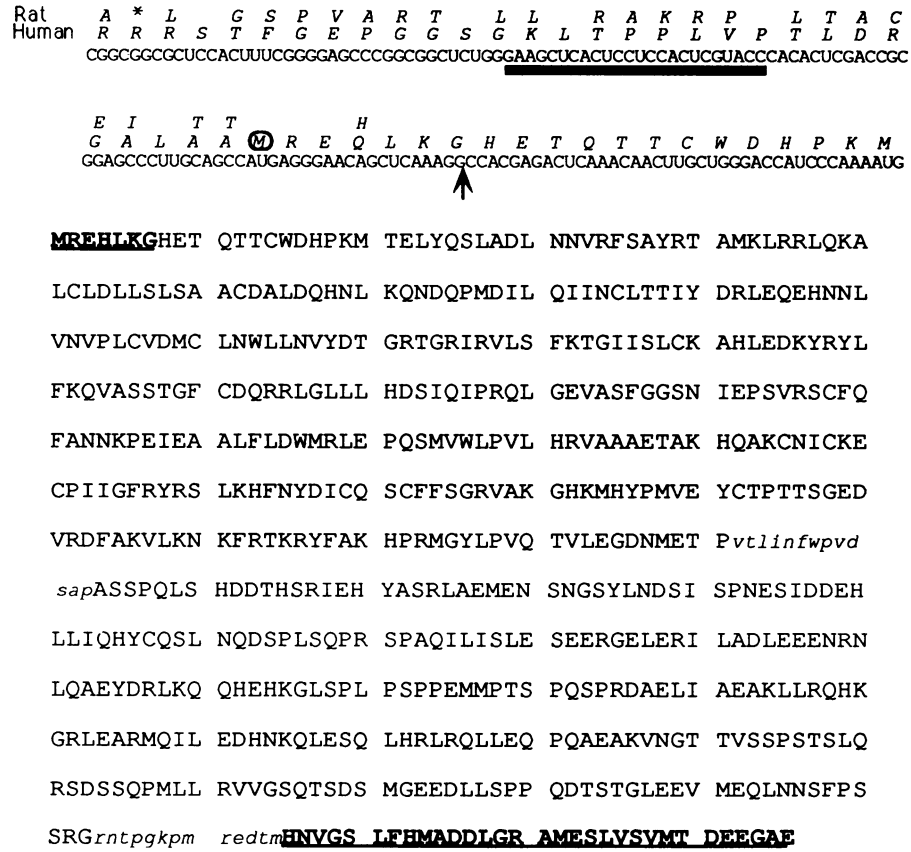

FIG. 1. (*Upper*) Deduced N terminus of the 70.8-kDa DMD gene product. The nucleotide sequence of the human 6.5-kb mRNA (lower line) is derived from a cDNA clone. The arrow indicates the location of an intron and the point of divergence between the sequence that is unique to the 6.5 kb mRNA and the sequence common to dystrophin mRNA and the 6.5-kb mRNAs. The deduced amino acid sequence is given above the cDNA in italics. Amino acids that are different in the rat sequence are indicated in the topmost line. The position of the stop codon in the mouse and rat amino acid sequence is indicated by the asterisk. The initiator methionine is encircled. The sequence used as a 5′ primer in PCR is underlined. The divergence between the rat and human sequence is due to divergence at the nucleotide level and to a frame shift. (*Lower*) Deduced amino acid sequence of the cloned coding region of the 6.5-kb mRNA. Amino acids that are not present in the published sequence of dystrophin are written in bold letters and underlined. The amino acids of dystrophin that are missing in the product caused by alternative splicing are written in italics.

exon (32 bp) changes the reading frame; the 13 C-terminal amino acids of dystrophin are replaced by 31 new amino acids in the protein product of the 6.5-kb mRNA (Fig. 1 *Lower*). Thus, the protein product of the 6.5-kb mRNA contains the C-terminal domain and the cysteine-rich domains of dystrophin (except for the modification described above) and only a very short additional peptide of 7 amino acids in the N terminus (Fig. 1 *Lower* and Fig. 2); this protein contains only 622 amino acids, and its calculated molecular weight is 70.8 kDa.

***In Vitro* Transcription and Translation of the Cloned Sequence.** *In vitro* transcription and translation of the protein encoded by the cDNA insert in plasmid HMD2.1 produced on a polyacrylamide/SDS gel a major protein band that migrates as a 77-kDa protein (Fig. 3 *Left*). The 10% discrepancy between the calculated and the apparent molecular weight of the protein is probably due to abnormal migration of the protein in the gel. We therefore tentatively refer to this polypeptide as a 71-kDa protein (Dp71). As expected (Fig. 2), the *in vitro* translation product was precipitated by the mAb MANDRA1, which is specific for an epitope in the C-terminal domain, but not by antibodies specific for the spectrin-like domain of dystrophin (Fig. 3 *Right*).

Dp71 Is Expressed in Many Tissues. To identify the polypeptide encoded by the 6.5-kb mRNA in cell extracts, protein extracts from various rat tissues and human hepatoma Hep G2 cells were analyzed in Western blots and stained with the mAb MANDRA1. This mAb reacts specifically with the X-linked dystrophin but not with the 400-kDa dystrophin-related protein, which is the product of a gene located on human chromosome 6 (21). MANDRA1 strongly stained dystrophin in muscle extracts (the 400-kDa band) (Fig. 4 *Upper*). In brain extracts, the antibody weakly stained dystrophin, but strongly stained a polypeptide comigrating with the *in vitro* 71-kDa cell-free translation product (Figs. 3 and 4 *Upper*). Only the Dp71 was detectable in liver extracts, while both dystrophin and Dp71 bands were present in extracts of stomach tissue, which is mostly smooth muscle. The tissue distribution of Dp71 kDa is very similar to that of the 6.5-kb mRNA as determined by Northern blots and RNase protection analysis (15, 23). The same results were obtained with another dystrophin-specific mAb (H-2A12) raised against a different epitope in the C-terminal domain of chicken dystrophin and with four additional mAbs that crossreacted with the dystrophin-related protein. Several dystrophin-specific antibodies, directed against unidentified epitopes in the C-terminal domain, did not react with Dp71 (data not shown). This may be attributed to differences in exon splicing between the mRNAs encoding the 400- and 71-kDa proteins described above.

As with the *in vitro* translation product, the cell-extracted protein did not react with mAb raised against epitopes in the spectrin-like domain (Fig. 4 *Lower*). Nor did it react with antibodies raised against epitopes in the last 10 amino acids of dystrophin, which are missing in the 6.5-kb translation product (not shown).

mdx mice do not synthesize dystrophin because of a point mutation in the *DMD* gene, resulting in a stop codon at

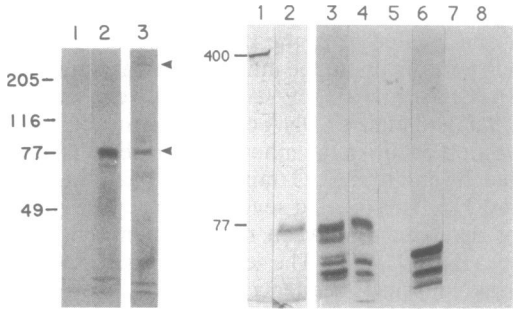

FIG. 3. (*Left*) *In vitro* transcription/translation of plasmid HMD2.1. *Hin*dIII-linearized plasmid HMD2.1 DNA (0.5 μg) was transcribed *in vitro* by using phage T7 RNA polymerase. The resulting RNA was translated in a rabbit reticulocyte lysate (Stratagene) with [^{35}S]methionine as the radioactive label. Cell-free translation products were fractionated on a 3–10% polyacrylamide/SDS gradient gel (18, 19) and blotted onto a nitrocellulose sheet. The blot was immunostained with mAb MANDRA1 and then autoradiographed. Lanes: 1, no RNA added; 2, ≈0.1 μg of RNA added to the reticulocyte lysate (lanes 1 and 2) represents the autograph of the translation product; and 3, immunostaining of 30 μg of rat brain protein extract electrophoresed on the same gel. The positions of size markers in kDa are indicated on the left. The black arrowheads indicate the migration positions of dystrophin and the Dp71. (*Right*) Immunoprecipitation of the cell-free product. Plasmid HMD-2.1 was *in vitro* transcribed and translated as described in *Left*. Cell-free products were immunoprecipitated with the antibodies specified by using rabbit anti-mouse immunoglobulin antibodies and protein A-Sepharose. The samples were fractionated on a 3–10% polyacrylamide/SDS gel, stained with mAb MANDRA1 (lanes 1 and 2), and autoradiographed (lanes 3–8). Lanes: 1, muscle extract; 2, brain extract that was mixed with the cell-free product shown in lane 3; 3, unprecipitated cell-free products; 4 and 5, cell-free product precipitated with MANDRA1 and MANDYS1, respectively; 6, translation product of ornithine decarboxylase mRNA; and 7 and 8, the same product precipitated with MANDRA1 or MANDYS1, respectively.

position 3185 of the dystrophin mRNA (25). As the 6.5-kb mRNA does not contain this region of the dystrophin mRNA, we anticipated that if the cellular Dp71 protein is a genuine product of the 6.5-kb mRNA, its production would not be affected by the mutation that abolishes the production of dystrophin in the *mdx* mice. Indeed, mAb MANDRA1 strongly stained dystrophin and Dp71 in extracts of the relevant tissues of normal mice, whereas in extracts of nonmuscle tissues of *mdx* mice, only the 71-kDa protein was detectable (not shown). These results rule out the possibility that Dp71 is a degradation or cleavage product of dystrophin.

As is the case with normal muscles, Dp71 was undetectable in the muscle of *mdx* mice. This rules out the possibility that the surprisingly mild course of the disease in these mice is due to replacement of dystrophin by Dp71.

DISCUSSION

Early studies suggested that the *DMD* gene is expressed specifically in the muscle and, to a much lesser extent, in the

FIG. 2. Schematic comparison of the domains encoded by the 14-kb dystrophin mRNA and the 6.5-kb mRNA. (*Upper*) Coding regions of the 14-kb brain and muscle dystrophin mRNAs. The regions coding for the various dystrophin domains are indicated on the bar and described below it. (*Lower*) Coding regions of the 6.5-kb mRNA. The divergence between the sequences common to the dystrophin mRNA and the 6.5-kb mRNA is at the 5′-end of the sequence coding for the cysteine rich-domain. The modifications created by alternative initiation and splicings are detailed in Fig. 1 *Lower*.

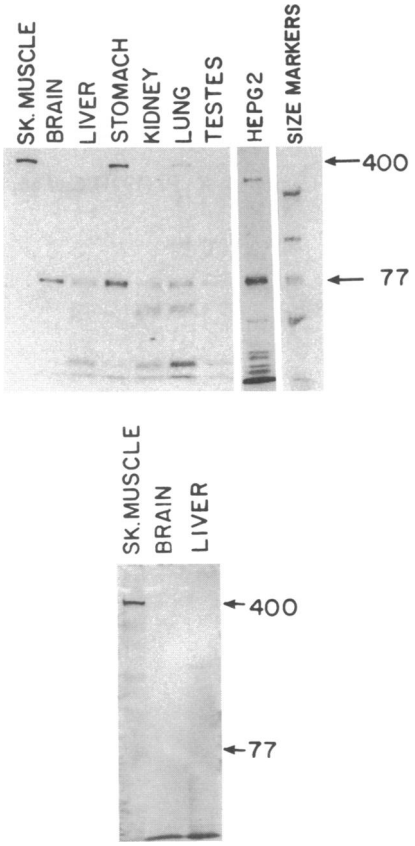

FIG. 4. Identification of the 71-kDa protein in tissue and cell extracts. (*Upper*) Protein samples (30–40 μg) isolated from the indicated rat tissues and Hep G2 cells were size-fractionated on a 3–10% polyacrylamide/SDS gel, blotted onto a nitrocellulose sheet, and immunostained with the mAb MANDRA1. The positions of dystrophin (400 kDa) and Dp71 (77 kDa) on the blot are indicated by arrows. Most of the other faint bands were also stained with the second antibody only. (*Lower*) Same as in *Upper*, but the first antibody used for the staining was MANDYS1, which reacts with an epitope in the spectrin-like domain of dystrophin.

brain. It now becomes evident that the *DMD* gene is expressed in many, and perhaps all, tissues and that this huge gene encodes at least one additional protein that differs greatly from the two known closely related dystrophin isoforms. Furthermore, this novel protein is the major *DMD* gene product detectable in the brain and other nonmuscle tissues. In some of them, the amount of this protein is comparable to the amount of dystrophin in the muscle. Its relatively high abundancy in brain cells is of special interest because it raises the possibility that it may be causally involved in some cases of mental retardation that are associated with DMD.

The function of this *DMD* gene product is still unknown. The predicted amino acid sequence shows that Dp71 lacks the spectrin-like domain that confers a rod-shaped structure to dystrophin and it also lacks the N-terminal domain, which is assumed to bind actin. However, it contains the C-terminal domain, which is the most conserved domain of dystrophin (26), and the cysteine-rich domain, which shares homology with the Ca^{2+}-binding domain of α-actinin. It has been shown that dystrophin forms a complex with glycoproteins, which are apparently responsible for its association with the cell membranes. It seems that this association is mediated by the C-terminal domain (27). Preliminary experiments indicate that Dp71 is associated with the cell-membrane fraction (our unpublished data). It is of interest to find out whether Dp71 forms complexes with the same or different glycoproteins. It should be noted that the last 13 amino acids of dystrophin, which are hydrophilic, are replaced by 31 amino acids that confer hydrophobicity to this region in Dp71. These 31 amino acids can form an α-helix, a part of which is amphipathic. This may result in a change in the function of this region. When one takes all of these data together, it is conceivable that the function of Dp71 is different from that of dystrophin.

As predicted from the size and structure of the 6.5-kb mRNA, Dp71 should not be affected by a great proportion of mutations in the *DMD* gene that abolish the production of dystrophin (as in the *mdx* mice case). A systematic comparison of the expression of Dp71 protein in DMD patients and the pathology of the disease in nonmuscle tissues might be very informative.

We thank Dr. K. Campbell for the antibodies directed against the C-terminal 10 amino acids of dystrophin. We gratefully acknowledge Ms. Dorit Zuk and Ms. Vivienne Laufer for editorial assistance. This work was supported by research grants from the Muscular Dystrophy Association of the U.S.A. and the Association Française contre les Myopathies, the Israel Academy of Sciences, the Israel Ministry of Health, the Leo and Julia Forchheimer Center for Molecular Genetics, and the Kekst Foundation. U.N. is the incumbent of the Elias Sourasky Professorial Chair.

1. Moser, H. (1984) *Hum. Genet.* **66,** 17–40.
2. Koenig, M., Monaco, A. R. & Kunkel, L. M. (1988) *Cell* **53,** 219–228.
3. Arahata, K., Ishiura, S., Ishiguro, T., Tsukahara, T., Suhara, Y., Eguchi, C., Ishihara, T., Nonaka, I., Ozawa, E. & Sugita, H. (1988) *Nature (London)* **333,** 861–863.
4. Sugita, H., Arahata, K., Ishiguro, T., Suhara, Y., Tsukahara, T., Ishiura, S., Eguchi, C., Nonaka, I. & Ozawa, E. (1988) *Proc. Japan Acad. Sci.* **64,** 37–39.
5. Watkins, S. C., Hoffman, E. P., Slayter, H. S. & Kunkel, L. M. (1988) *Nature (London)* **333,** 863–866.
6. Zubrzycka-Gaarn, E. E., Bulman, D. E., Karpat, G., Burghes, A. H. M., Belfall, B., Klamut, H. J., Talbot, J., Hodges, R. S., Ray, P. N. & Warton, R. G. (1988) *Nature (London)* **333,** 466–469.
7. Nudel, U., Robzyk, K. & Yaffe, D. (1988) *Nature (London)* **331,** 635–638.
8. Nudel, U., Zuk, D., Einat, P., Zeelon, E., Levy, Z., Neuman, S. & Yaffe, D. (1989) *Nature (London)* **337,** 76–78.
9. Yaffe, D., Zuk, D., Einat, P., Shinar, D., Levy, Z., Neuman, S., Fuchs, O. & Nudel, U. (1989) in *Cellular and Molecular Biology of Muscle Development*, eds. Kedes, L. H. & Stockdale, F. E. (Liss, New York), pp. 963–971.
10. Barnea, E., Zuk, D., Simantov, R., Nudel, U. & Yaffe, D. (1990) *Neuron* **5,** 881–888.
11. Feener, C. A., Koenig, M. & Kunkel, L. M. (1989) *Nature (London)* **338,** 509–511.
12. Makover, A., Zuk, D., Breakstone, J., Yaffe, D. & Nudel, U. (1991) *Neuromusc. Disorders* **1,** 39–45.
13. Boyce, F. M., Beggs, A. H., Feener, C. & Kunkel, L. M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1276–1280.
14. den Dunnen, J. T., Casula, L., Makover, A., Bakker, B., Yaffe, D., Nudel, U. & van Ommen, G. J. B. (1991) *Neuron Dis.*, in press.
15. Bar, S., Barnea, E., Levy, Z., Neuman, S., Yaffe, D. & Nudel, U. (1990) *Biochem. J.* **272,** 557–560.
16. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
17. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
18. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
19. Pons, F., Augier, N., Leger, J., Robert, A., Tome, F. S. M., Fardeau, M., Voit, T., Nicholson, L. V. B., Mornet, D. & Leger, J. J. (1991) *FEBS Lett.* **282,** 161–165.
20. Ellis, J. M., Man, N. T., Morris, G. E., Ginjaar, I. B., Moorman, A. F. M. & van Ommen, G. J. B. (1990) *Lancet* **336,** 881–882.

21. Love, D. R., Hull, D. P., Dickson, G., Speerr, N. K., Bryth, B. C., Marsden, R. S., Walsh, P. S., Edwards, Y. H. & Davis, K. E. (1989) *Nature (London)* **339,** 55–58.
22. Nguyen, thi Man, Cartwright, A. J., Morris, G. E., Love, D. R., Bloomfield, J. R. & Davies, K. E. (1990) *FEBS Lett.* **262,** 237–240.
23. Rapaport, D., Lederfein, D., Den Dunnen, J. T., Grootscholten, P. M., van Ommen, G. J. B., Fuchs, O., Nudel, U. & Yaffe, D. (1992) *Differentiation* **49,** 187–194.
24. Kozak, M. (1986) *Cell* **44,** 283–292.
25. Sicinski, P., Geng, Y., Ryder-Cook, A. S., Barnard, E. A., Darlison, M. G. & Barnard, P. J. (1989) *Science* **244,** 1578–1580.
26. Lemaire, C., Heilig, R. & Mandel, J. L. (1988) *EMBO J.* **7,** 4157–4162.
27. Ervasti, J. M. & Campbell, K. P. (1991) *Cell* **66,** 1121–1131.